# Predicting High-Risk Behaviors in Veterans With Posttraumatic Stress Disorder

*Tamara L. Hartl, PhD,* Craig Rosen, PhD,†‡ Kent Drescher, PhD,† Tina T. Lee, MD,† and Fred Gusman, MSW†*

**Abstract:** The present study sought to identify posttraumatic stress disorder (PTSD) patients at high risk for negative behavioral outcomes (violence, suicide attempts, and substance use). The Mississippi Scale for Combat-Related PTSD, the Beck Depression Inventory, and demographic and behavioral data from 409 male combat veterans who completed a VA residential rehabilitation program for PTSD were analyzed using signal detection methods (receiver operating characteristics). A validation sample ($N$ = 221) was then used to test interactions identified in the signal detection analyses. The best predictors of behaviors at follow-up were those same behaviors shortly before intake, followed by depressive and PTSD symptoms. However, for each of the models other than that for hard drug use, cutoffs determined at the symptom level did not lend themselves to replication. Recent high-risk behaviors, rather than patients' history, appear to be more predictive of high-risk behaviors postdischarge.

**Key Words:** PTSD, veterans, violence, suicide, substance use.

(*J Nerv Ment Dis* 2005;193: 464–472)

Posttraumatic stress disorder (PTSD) is a life-altering and complex psychological condition estimated to affect 30% of Vietnam theater veterans over their lifetime (Kulka et al., 1990). Despite recent increases in the number of treatment programs aimed at addressing the myriad of problems associated with the disorder (Chief Medical Director's Special Committee on PTSD, 1991), high-risk behaviors such as suicide attempts, violence, and substance use persist among PTSD veterans even after treatment (Rosenheck and Fontana, 2001).

For example, several studies indicate that veterans with PTSD endorse greater rates of interpersonal violence compared with veterans without PTSD (Beckham et al., 1997; Kulka et al., 1990). Greater risk of violence in PTSD veterans has been shown to be associated with more severe hyperarousal symptoms and drinking quantity (Savarese et al., 2001), lower socioeconomic status and PTSD severity (Beckham et al., 1997), and participation in war zone violence (Hiley-Young et al., 1995).

In addition to the greater risk of violence associated with PTSD in veterans, there is substantial evidence that a diagnosis of PTSD in Vietnam era veterans amplifies the overall risk of attempting or committing suicide (Bonin et al., 2000; Fontana and Rosenheck, 1995).

Substance abuse is also common among PTSD veterans, even after completion of PTSD treatment programs. Perconte and Griger (1991) assessed combat veterans with PTSD who completed a partial hospitalization program for PTSD. Compared with patients who relapsed, veterans who improved in treatment and maintained positive gains at 18-month follow-up endorsed a lower weekly alcohol intake at the time of intake and discharge.

On average, PTSD symptoms do not seem to be ameliorated by inpatient or outpatient treatment programs (Fontana and Rosenheck, 1996; Hammarbarg and Silver, 1994; Johnson et al., 1996). Therefore, relapse in this population might better be defined in terms of high-risk behaviors rather than symptom reoccurrence or exacerbation. Because of the chronicity of PTSD and its association with high-risk behaviors, there is a great need for predicting which patients are at greater risk for relapse and serious harm. Better prediction of outcomes could lead to specialized interventions that result in a decrease in rehospitalization rates and a substantial enhancement to veterans' quality of life.

The epidemiological and clinical studies discussed here provide ample evidence that veterans with chronic PTSD can

*Health Services Research and Development, VA Palo Alto Health Care System, Palo Alto, California; †National Center for PTSD, VA Palo Alto Health Care System, Palo Alto, California; and ‡Stanford University School of Medicine, Stanford, California.

Supported in part by the VA National Center for Posttraumatic Stress Disorder, VA Health Services Research and Development Service, VA Office of Academic Affiliations, VA Mental Health Strategic Health Group, and VA Palo Alto Health Care System.

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Send reprint requests to Tamara L. Hartl, PhD, 116 B, VA Palo Alto Health Care System, 3801 Miranda Ave., Palo Alto, CA 94304.



ISSN: 0022-3018/05/19307-0464
DOI: 10.1097/01.nmd.0000168238.13252.b3

be at risk for relapse with respect to a number of high-risk behaviors, including attempted or completed suicide, violence, and misuse of alcohol and drugs.

However, existing research provides little guidance for determining which individual PTSD patients are at greatest risk. There are at least three reasons for this. First, prediction of individual behavior is notoriously difficult, particularly when the behavior has a relatively low base rate (McNeil and Binder, 1997). For an infrequent behavior such as suicide, even a very good model will generate a high proportion of false positives to true positives. Second, patient subpopulations that are at the highest risk may be relatively homogenous in terms of risk factors. For example, it is not difficult to predict that psychiatric inpatients are at greater risk for violence than are psychiatric outpatients. However, it is more difficult to predict which inpatients are at highest risk, because several risk factors (such as past history of violence, history of substance misuse, and poor social support) may be nearly universal in this patient population. The final barrier is statistical. Conventional correlational analyses and logistic regression models reported in most research are primarily designed to identify which variables are associated with risk for an outcome, rather than which individuals are at greatest risk (Kiernan et al., 2001). Moreover, such analyses do not provide clinically useful cut-points or decision rules that can be used to categorize patients.

## CURRENT STUDY

The aim of the present study is to assess patient characteristics that predict relapse to high-risk behaviors following discharge from a residential rehabilitation program for PTSD. To evaluate patient characteristics that predict greater risk of relapse, we used signal detection methods that are commonly used in medical decision making to evaluate performance on diagnostic tests (Kraemer, 1992). Receiver operating characteristics (ROC) is a type of signal detection analysis that has specific advantage over logistic regression for the purposes of identifying which patients are at risk. Specifically, ROC is a person-centered approach that identifies individuals at high or low risk based on their scores on predictor variables. In contrast, logistic regression is variable-centered and identifies variables that are related to outcomes across individuals. The advantage of using ROC analysis for this investigation is that it yields an empirically derived tool for identifying high-risk patients and thus can inform clinicians which patients are most in need of follow-up support services. Other advantages of signal detection methods include the ability to designate a criterion for favoring detection of false positives over false negatives, reservation of power amid multicollinear predictors, and the designation of empirically constructed higher-order interactions (Kiernan et al., 2001). The primary outcome variables of interest for this study included whether patients (a) attempt suicide, (b) engage in violent behaviors, (c) misuse alcohol, or (d) misuse cocaine, amphetamines, or opiates within 4 months of discharging from the controlled environment of residential treatment.

## METHODS

### Sample

The sample consisted of 630 male veterans with a primary PTSD diagnosis who consecutively entered the residential rehabilitation program for PTSD at the VA Palo Alto Health Care System between July 1994 and December 2000. The residential rehabilitation program is a highly structured treatment in which every patient participates in various skills classes (e.g., conflict resolution, anger management, effective communication), medication management, relapse prevention, and recreation. Patients are referred to the program by medical and mental health staff from health care facilities largely concentrated in the Pacific Northwest.

For patients with multiple admissions during this time period ($N = 85$), the earliest admission was retained. Length of stay for the entire sample ranged from 2 to 286 days, with an average of 71.6 days ($SD = 35.29$). Although there is a large disparity in length of stay, the program's standard length of treatment was 90 days (before 1996) and then 60 days (1996 and on). The mean age of the sample was 50.56 ($SD = 4.55$; range = 25.79–76.26). Ethnicity was reported as 60% Caucasian, 13% Hispanic/Latin American, 13% African American, 6% mixed ethnicity, 5% Native American, 1% Asian/Pacific Islander, and 1% other. One percent of the responses was missing for ethnicity. Ninety percent of the men were Vietnam era veterans, 86% of whom had served in combat.

### Measures

The data set was constructed to reflect information that is typically collected at the outset of most treatment programs. Thus, any decision-making tools that might result from this investigation might be testable in, and perhaps generalizable to, other programs.

#### Demographic and Military Service Variables

Variables from demographic and background history questionnaires administered at intake included age, ethnicity, education, marital status, and history of incarceration. Nearly 72% of the sample had a history of incarceration (Table 1), which speaks to the severity of impairment in this patient population. Variables related to patients' war zone trauma exposure included whether patients served in a war zone, witnessed incoming fire, witnessed or participated in atrocities, or had been held a prisoner of war. Treatment-related variables included number of prior admissions to the same rehabilitation program and whether patients dropped out of

TABLE 1. Sample Characteristics at Intake in a Sample of 630 Male Veterans

| Variable | Mean | *SD* | Range |
|---|---|---|---|
| Education (y) | 13.30 | 1.89 | 6–20 |
| Mississippi Sale scores | 138.02 | 15.74 | 88–175 |
| BDI scores | 32.26 | 10.56 | 0–59 |
| Number of admissions | 1.17 | .52 | 1–8 |
| Number of types of violence | 1.14 | 1.09 | 0–3 |
| Number of hard drug SCID Dx's | .65 | .93 | 0–5 |
| Number of days drank alcohol prior to intake | 3.2 | 8.08 | 0–30 |
| Number of days intoxicated prior to intake | 2.22 | 7.01 | 0–30 |
| Number of days used cannabis prior to intake | 1.67 | 6.27 | 0–30 |
| Number of days used hard drugs prior to intake (aggregate) | 1.68 | 7.69 | 0–90 |
| Variable | Percentages | | |
| Marital status | 36.5% married | | |
| Served in a war zone | 98.4% | | |
| Witnessed incoming fire | 97.6% | | |
| Prisoner of war | 1.6% | | |
| Witnessed or participated in atrocities | 60.8% | | |
| Suicide attempt 4 mo prior to intake | 8.1% | | |
| Suicide attempt ever | 50.2% | | |
| Had been incarcerated | 71.6% | | |
| Left treatment early | 10.3% | | |
| SCID Dx of alcohol abuse or dependence | 72% | | |
| SCID Dx of cannabis abuse or dependence | 27% | | |
| SCID Dx of hard drug abuse or dependence | 39% | | |

the program early, defined as staying fewer than 30 days when the program's typical length of stay was 60 days and staying fewer than 45 days when the program's average length of stay was 90 days.

## Behaviors

Violent behavior and suicide attempts were assessed with items from the Northeast Program Evaluation Center survey used in ongoing program evaluation of VA PTSD patients (Fontana and Rosenheck, 1997; McFall et al., 1999). Three violence items inquired whether in the last 4 months patients had threatened someone with physical violence, had a physical fight with someone, or threatened someone with a weapon. Two additional items asked whether patients had attempted suicide in their lifetimes and/or in the last 4 months. Half of the sample had attempted suicide at some point (Table 1). Substance use was assessed with several Northeast Program Evaluation Center items drawn from the Addiction Severity Index (McLellan et al., 1992). Questions about drug and alcohol use specified that respondents should report their use in the 30 days prior to a possible qualifying period (typically 30 days) of no substance use for the program. Questions included the number of days the patient drank alcohol, drank alcohol to the point of feeling drunk or intoxicated, and used cannabis, cocaine, amphetamines, or opiates. The last three items, reflecting the most commonly used hard drugs in this population, were summed into a single measure of hard drug use. For example, if a participant used cocaine on 20 days, amphetamines on 15 days, and opiates on 0 days, his total hard drug use score would equal 35.

## Symptoms

Posttraumatic stress disorder symptoms were assessed with the Mississippi Scale for Combat-Related PTSD (Keane et al., 1988). The Mississippi Scale is a reliable and well-validated 35-item measure of PTSD symptoms in combat veterans (Keane et al., 1987; Kulka et al., 1990). Total scores from the Beck Depression Inventory-II (BDI-II; Beck et al., 1996) were used to assess depressive symptoms.

Due to multicollinearity between patient symptom scores from intake and discharge assessments, we included only intake scores as predictors of postdischarge outcome. Intake scores were selected because these data were more complete, and if predictive, could be most helpful in informing treatment planning based on patients' profiles at intake. Additional post hoc analyses (detailed below) tested whether using discharge rather than intake symptom scores improved our prediction of postdischarge outcomes.

## Diagnoses

Structured Clinical Interview for DSM-IV Disorders (First et al., 1995) diagnoses were included for alcohol, cannabis, and hard drugs. The last was a composite variable reflecting a diagnosis of cocaine, amphetamine, or opioid dependence. These three variables were coded in order of increasing severity from "no diagnosis" to "dependence in early or partial remission." Number of hard drug diagnoses (from 0–5: cocaine, amphetamines, opioids, sedatives, and hallucinogens) was included as another potential predictor. Major depression, bipolar disorder, and anxiety disorder diagnoses were not included because preliminary bivariate

analyses showed no relationship between these variables and the outcomes of interest.

### Outcomes

Four dependent variables (DVs) were selected to reflect relapse to high-risk behaviors. All DVs were binary to enable ROC analyses to be applied to the data set. Violence was defined as engaging in violent behavior (physical fight, threatening someone, or threatening with a weapon) in the 4 months after discharge. Suicide was defined as attempting suicide within the 4 months following discharge. Alcohol misuse was defined as "drinking to the point where you felt drunk or intoxicated; that is, had three or more drinks in one sitting" on 2 or more days in the month prior to follow-up. Intoxication two or more times in the past 30 days, as opposed to 1 or more days, was chosen to discriminate from drinking to intoxication on only 1 day (the most frequently endorsed response). Hard drug use was defined as any use of cocaine, amphetamines, or opioids in the month prior to follow-up. Whereas the questions about violence and suicide attempts assessed these behaviors in the past 4 months, questions about substance abuse applied only to the 30 days prior to completing the follow-up questionnaire.

## Data Analytic Plan

Software specifically designed to compute ROCs (Kraemer, 1992) was used in an effort to identify interactions between predictors. ROC is a signal detection method that entails the use of a cutoff point for each predictor to procure the best differentiation on the dependent variable in question based on the assigned sensitivity/specificity weight. For this investigation, the weight was set at .70 to favor sensitivity and avoid false negatives. In this type of analysis, the sample is divided into two subsamples at the optimal cutpoint on each variable that predicts the specified clinical criterion (the DV). The process is then repeated in an iterative manner across all remaining predictors in the two subsamples to detect the next best predictor of the criterion. The process continues until too few individuals for analysis remain in the subgroup (e.g., $n < 10$).

Receiver operating characteristics analyses were first conducted on a random selection of two thirds (409 patients) of the sample to establish prediction models for each outcome. The remaining one third of the patients ($N = 221$) were used as a replication sample to test the fit of the models developed in the exploratory analyses.

# RESULTS

## Descriptive Statistics

In the 30 days prior to a prequalifying period of abstinence, 11.6% of the sample reported having drunk to the point of intoxication on 2 or more days, and 8.1% had used drugs. In the 4 months prior to intake, 61.9% of the sample reported having committed one or more violent acts, and 8.1% reported having attempted suicide. Reports of these behaviors at the 4-month follow-up assessment indicate a decrease in both violence (49% of the sample committed one or more violent acts) and suicide attempts (4.9% of the sample) over the preceding 4 months. However, in the 30 days preceding follow-up, drinking to intoxication and drug use both increased slightly relative to preintake reports (14.3% and 9.5%, respectively). This is likely a reflection of individuals returning to drinking after a required period of abstinence for program admission.

## Signal Detection: Exploratory Sample

Figures 1 to 4 depict the ROC models for the four behavioral DVs assessed at follow-up. Each figure depicts a decision tree type of graph to delineate partitioning off of subgroups of patients based on cutoff scores for the significant predictors in each model. The base rate of the behavior being predicted in the exploratory samples was used as the cut-point to discriminate high-risk versus low-risk groups for each model, with groups scoring at or above the base rate considered high-risk. Sensitivity indices were calculated by summing the high-risk patients in each of the subgroups at the end of a tree branch and then dividing the number of actual relapses (true positives) by the total number of patients that relapsed. The specificity index equaled the number of true negatives divided by the total number of patients who did not score positively on the outcome variable.

### Replication

The predictive validity of the four ROC models was evaluated by using the cut-point determined in the explor-

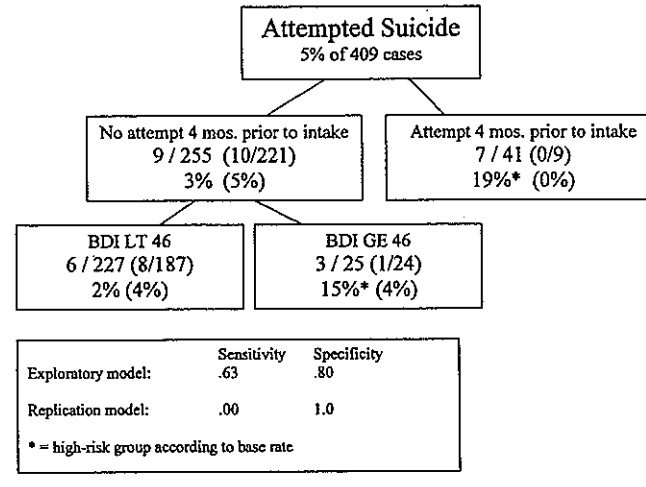
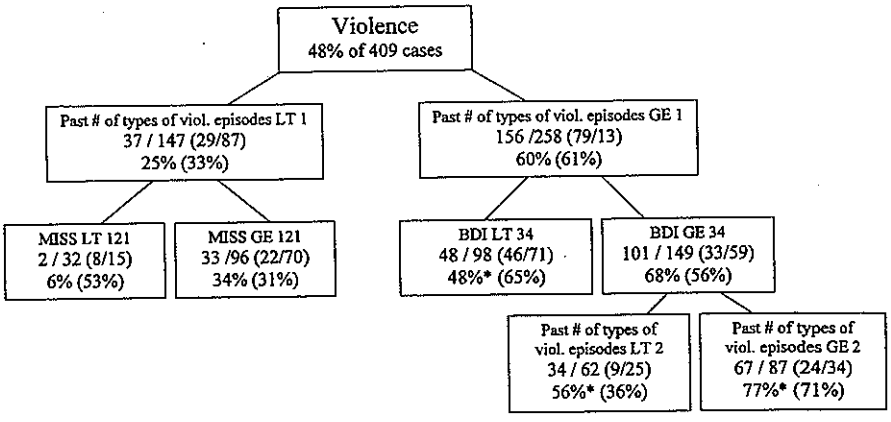


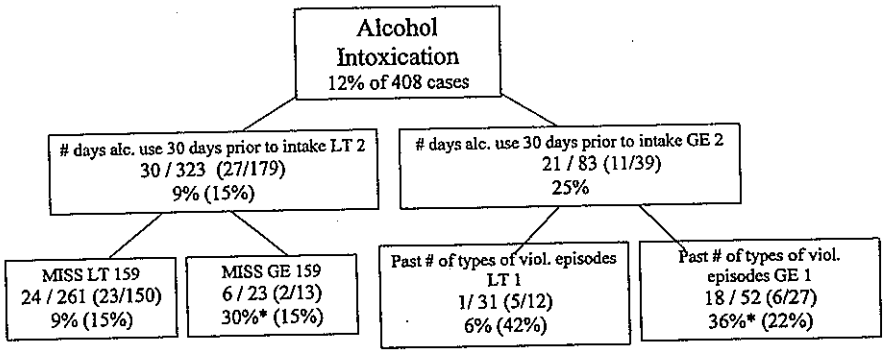

**FIGURE 1.** Prediction model for suicide. Figures are for the exploratory sample ($N = 408$); figures in parentheses are for the replication sample ($N = 221$).




| | Sensitivity | Specificity |
|---|---|---|
| Exploratory model: | .81 | .49 |
| Replication model: | .72 | .60 |

* = high-risk group according to base rate

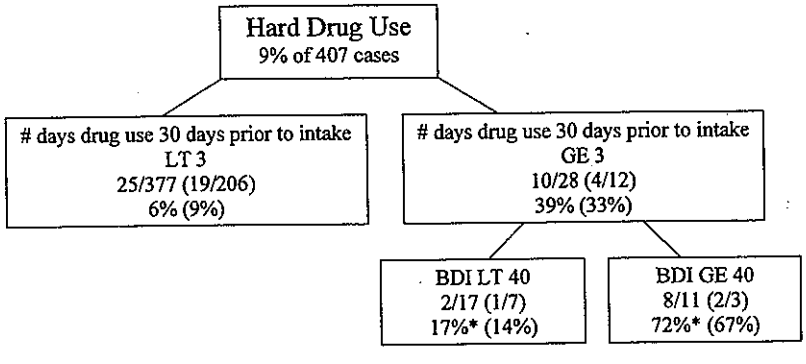
**FIGURE 2.** Prediction model for violence. Figures are for the exploratory sample ($N$ = 408); figures in parentheses are for the replication sample ($N$ = 221).




| | Sensitivity | Specificity |
|---|---|---|
| Exploratory model: | .49 | .84 |
| Replication model: | 1.0 | .00 |

* = high-risk group according to base rate

**FIGURE 3.** Prediction model for alcohol intoxication (three or more times in the past month). Figures are for the exploratory sample ($N$ = 408); figures in parentheses are for the replication sample ($N$ = 221).




| | Sensitivity | Specificity |
|---|---|---|
| Exploratory model: | .29 | .95 |
| Replication model: | .17 | .96 |

* = high-risk group according to base rate

**FIGURE 4.** Prediction model for hard drug use in the past month. Figures are for the exploratory sample ($N$ = 408); figures in parentheses are for the replication sample ($N$ = 221).

atory sample to categorize subjects in the replication sample. The percentages of positive scores on the DVs within the replication sample at each branch or cutpoint are shown in parentheses in Figures 1 to 4. Overall sensitivity and specificity for each model were also calculated in the replication sample.

## Suicide

### Exploratory Sample

Signal detection methods identified predictors that distinguished patients more and less likely to attempt suicide after discharge. In Figure 1, two high-risk subgroups were identified. The single best predictor of a suicide attempt after discharge was having attempted suicide in the 4 months prior to intake ($\chi^2[1,296] = 15.03$; $p < 0.001$). Among those who had not attempted suicide shortly before intake (left side of Fig. 1), the next optimal predictor was patients' BDI scores: 15% of patients whose BDI scores were greater than or equal to 46 had attempted suicide at follow-up, as opposed to 2% whose BDI scores were less than 46 ($\chi^2[1,252] = 10.54$; $p < 0.001$). Sensitivity for this model was calculated at .63, with a specificity of .80.

### Replication

For the suicide model, the first cutoff determined by prior attempt within the previous 4 months did not replicate well. Compared with 19% of the original sample who both had a recent attempt and attempted in the 4 months postdischarge, 0% of the replication sample who had a recent prior attempt attempted postdischarge. Sensitivity and specificity for the model in the replication sample were .11 and .84, respectively.

## Violence

### Exploratory Sample

For violence (Fig. 2), the single best predictor of committing a violent act at follow-up was having committed one or more violent acts prior to intake ($\chi^2[1,405] = 45.96$; $p < 0.001$). Among those who had committed one or more violent acts, the next best predictor was BDI scores ($\chi^2[1,247] = 9.40$; $p < 0.001$), with patients scoring 34 or greater comprising a higher risk group (68% vs. 48%). There were no further distinctions for patients scoring less than 34 on the BDI. The number of past violent acts reemerged as the best predictor ($\chi^2[1,149] = 7.09$; $p < 0.001$) among the subgroup of patients with higher BDI scores, such that patients who committed two or more violent acts represented the highest risk group in the model (77%).

Among patients who reported no violence prior to intake, the best predictor of postdischarge violence was Mississippi PTSD scores ($\chi^2[1,128] = 9.56$; $p < 0.001$), with patients scoring greater than or equal to 121 comprising another high-risk group (34%). In comparison, patients scoring less than 121 on the Mississippi constituted a low-risk group (6% likelihood of being positive for follow-up violence). Overall sensitivity of the model was .81, while specificity was .49.

### Replication

The first cutoff in the violence model replicated well, with similar percentages of patients reporting violence postdischarge among the first split's two subgroups, defined by past number of types of violent episodes. However, at the second split on the left side of the model, determined by a Mississippi cutoff of 121, replication was poor in that 53% of individuals with Mississippi scores lower than 121 reported violence at follow-up, compared with only 6% of the original sample. Similarly, the subgroup that had BDI scores greater than or equal to 34 and fewer than two past violent episodes comprised a high-risk group in the exploratory analysis but a low-risk group upon replication. Sensitivity for the replication model was .72, while specificity was .60.

## Intoxication

### Exploratory Sample

The optimal predictor for intoxication at follow-up was having been intoxicated 2 or more days during the 30-day period prior to the month of required abstinence before intake ($\chi^2[1,406] = 14.58$; $p < 0.001$; Fig. 3). Among those who had been intoxicated more than two times prior to intake, the next distinction was the commission of violent acts prior to intake, with patients who committed one or more violent act at greater risk ($\chi^2[1,83] = 9.30$; $p < 0.001$). An additional split for the subgroup of patients who had not committed violence was age, but the marginal means (cell sizes) were too small for this to be reliable, and thus it was excluded from the overall model. On the left side of the figure, among patients who had been intoxicated fewer than two times in the month prior to intake, the best predictor of violence was Mississippi scores ($\chi^2[1,284] = 9.81$; $p < 0.001$), with patients scoring greater than or equal to 159 representing a high-risk group (30% vs. 9%). In the exploratory sample, sensitivity for this model was .49, with a specificity of .84.

### Replication

For alcohol intoxication, there were several replication problems. The first split, determined by number of days within the last 30 that alcohol was used, yielded results similar to the original model, with 28% of the sample getting intoxicated 2 or more days at follow-up if they had used prior to intake 2 or more days, and 15% of the sample getting intoxicated at follow-up if they had drank alcohol fewer than 2 days prior to intake. However, there were large percentage differences on the outcome variable for patients who had used alcohol 2 or more days in the 30 days prior to intake and who had no past violent episodes (6% in original model vs. 42%

in the replication sample). Furthermore, Mississippi scores among those who had fewer than 2 days of intoxication prior to intake determined different outcomes between the original and replication samples. Sensitivity was 100%, while specificity was 0%.

## Hard Drug Use

### Exploratory Sample

The best predictor of drug use at follow-up was the number of days in the 30 days prior to the month of required abstinence before intake that patients used drugs, with those using on 3 or more days having significantly higher risk ($\chi^2[1,405] = 45.96$; $p < 0.001$; Fig. 4). Among those who had used drugs 3 or more days, an additional significant predictor was BDI scores, with patients scoring 40 or above posing the greatest risk for drug use at follow-up ($\chi^2[1,28] = 8.50$; $p < 0.001$). Sensitivity of this model was .29, with a specificity of .95.

### Replication

The model for hard drug use replicated well, with no more than 6% point differences among the two splits in the model. For this model, sensitivity was .17, and specificity was .96.

## Post Hoc Tests of Alternate Models

While the low base rate for attempted suicide at follow-up likely accounted for the replication problems in this model, the models for intoxication and violence included cut-off scores on the Mississippi PTSD scale that were near the tails (extremely high or low), and thus might be unreliable and unlikely to replicate in a new sample. In an attempt to build more replicable models for violence and intoxication, we replaced Mississippi Scale scores with BDI scores, as $\kappa$ scores for the BDI were generally similar to those of the Mississippi Scale (Table 2). However, when BDI scores were retained as the only symptom measure of distress and Mississippi scores were excluded from the model, the problems with replication remained.

Additional ROC models were run to determine whether better replication would ensue as a result of using discharge scores on the Mississippi and BDI instead of intake scores. However, none of the models, except for the violence model, were better replicated by this substitution. The violence

**TABLE 2.** Optimal Cutpoints and $\kappa$ Values at the First Cut for Variables in the ROC Analysis

| | Suicide | | Violence | | Alcohol | | Hard Drugs | |
|---|---|---|---|---|---|---|---|---|
| Variable | Best cutoff | $\kappa$ | Best cutoff | $\kappa$ | Best cutoff | $\kappa$ | Best cutoff | $\kappa$ |
| Age | 45 | .059 | 51 | .135 | 51 | .064 | 48 | .075 |
| Atrocities | 2 | .111 | 1 | .076 | 2 | .121 | 2 | .055 |
| Education | 17 | .048 | 14 | .097 | 15 | .069 | 14 | .059 |
| MISS intake | 161 | .212 | 134 | .230 | 145 | .106 | 153 | .075 |
| BDI intake | 46 | .203 | 32 | .223 | 42 | .128 | 40 | .148 |
| Number admits | 2 | .061 | 2 | .046 | 2 | .036 | 2 | .014 |
| Suicide attempt in past 4 mo | 1 | .246 | 1 | .028 | 1 | .033 | 1 | .040 |
| Suicide attempt ever | 1 | .056 | 1 | .062 | 1 | .034 | 1 | .001 |
| Incarceration | 1 | .009 | 1 | .057 | 1 | .031 | 1 | .032 |
| Marital status | 1 | .003 | 1 | .009 | 1 | .018 | 1 | .023 |
| Ethnicity | 1 | .014 | 1 | .006 | 1 | .047 | 1 | .046 |
| Dropout | 1 | .180 | 1 | .018 | 1 | .012 | 1 | .009 |
| Improvement | 1 | .058 | 0 | .022 | 0 | .074 | 0 | .045 |
| Violence | 2 | .028 | 1 | .366 | 1 | .086 | 1 | .044 |
| Alcohol Dx | 1 | .043 | 3 | .052 | 2 | .091 | 3 | .059 |
| Cannabis Dx | 1 | .018 | 2 | .082 | 2 | .032 | 1 | .032 |
| Hard drug Dx | 1 | .030 | 2 | .093 | 2 | .041 | 1 | .085 |
| Number of hard drug Dx | 1 | .043 | 2 | .043 | 2 | .117 | 2 | .150 |
| Alcohol use last 30 | 1 | .045 | 3 | .048 | 2 | .205 | 2 | .148 |
| Drunk last 30 | 1 | .025 | 28 | .031 | 3 | .180 | 2 | .024 |
| Cannabis use last 30 | 2 | .031 | 30 | .021 | 2 | .132 | 28 | .147 |
| Drug use last 30 | 1 | .091 | 25 | .026 | 17 | .115 | 3 | .266 |

model replicated well primarily because past violence was the only significant predictor that emerged.

## DISCUSSION

Our goal in this study was to identify which patients in a residential rehabilitation program for PTSD constituted high-risk subgroups likely to re-engage in problematic behaviors after discharge. We had mixed success.

For drug use, we were successful in reliably differentiating groups of patients at very high risk for drug use after discharge (i.e., those with recent drug use and very severe depression symptoms at intake) from those at moderate risk (those with recent drug use and BDI scores below 40) and relatively low risk (those without recent drug use). The exploratory model replicated well but did not achieve good sensitivity (.29 in the exploratory model and .17 in the replication model). The models were highly specific, however (exploratory model = .95; replication model = .96). Although recent drug use and depression are intuitive and already well-documented risk factors, the model developed here provided specific cut-points and decision rules that could be used to target follow-up services to those PTSD patients who are most at risk.

Our efforts to predict relapse to violence and alcohol misuse were less successful. On the one hand, recent behavior (violence or drinking to intoxication) prior to intake consistently predicted risk of relapse in both models. The additive effects of symptom severity, however, were unclear. More severe PTSD and/or depression symptoms were associated with greater risk of violence or alcohol misuse in the exploratory sample but failed to predict these outcomes reliably in the replication sample. Moreover, we were entirely unable to reliably differentiate patients at higher and lower risk for attempted suicide after discharge. None of the predictors identified in the exploratory sample (not even recent suicide attempts prior to intake) replicated in the second sample. The difficulty of predicting suicide in this population is likely a function of both the low base rate (in statistical terms, although not in epidemiological terms) of suicide attempts within any month period in contrast with the very high proportion of patients at potential risk for suicide (roughly half had a prior history of suicide attempts).

Considering the results from all four of these models leads us to several conclusions. Our first conclusion is that statistical replication is essential in testing the predictive value of any given ROC model. ROC analysis, like stepwise regression or factor analysis, is an empirically driven descriptive technique in which models are built by maximizing their fit with a given data set. The robustness of the resultant models can be tested only through replication. Without replication, we would have had much more confidence than is warranted in our ability to use specific cut-points on PTSD or depression symptom measures to identify patients at higher or lower risk for violence and alcohol misuse.

Our second conclusion is that recent behavior is the strongest predictor of future behavior. This is hardly surprising. However, it was somewhat surprising that lifetime history of behavior, at least in the case of lifetime history of suicide attempts and Structured Clinical Interview for DSM-IV Disorders substance use diagnoses, had little predictive value and may be much less informative in treatment planning. Demographic information and severity of military trauma also were not predictive of relapse in this sample. These findings highlight the importance of assessing and explicitly integrating recent problem behaviors into treatment planning. For example, typical chart notes in a mental health clinic may indicate that a patient carries a diagnosis of substance use in remission but may not provide information on last use and may not flag patients in early recovery as requiring extra follow-up.

Homogeneity may reduce the predictive effects of symptoms. Although there was considerable statistical variability within our sample in level of symptoms, all patients were strongly symptomatic. For example, the mean Mississippi PTSD Scale score in our sample was over a standard deviation higher than in a validation sample drawn from an outpatient VA Vet Center (Keane et al., 1988). The predictive utility of PTSD symptom scores may be reduced when the scores are so severe. In essence, the symptom severity in our sample may have been so severe as to have lost some of its meaningfulness in terms of discriminating patient outcomes.

In chronic patient populations, symptoms have a large component that is stable over time. For example, in our sample, the correlation between PTSD symptoms at intake and at discharge was .57. However, symptoms also fluctuate markedly over time. It may be the fluctuations in symptoms that are most predictive of high-risk behaviors, i.e., patients accustomed to a certain level of chronic distress, may become more likely to engage in high-risk behaviors at times when their symptoms become even worse. Thus, predicting risk of relapse may require more proximal measurement of symptoms than those obtained at intake, or even at discharge. The possible predictive value of changes in acute symptoms also suggests the importance of community-based interventions and frequent postdischarge follow-ups to help patients better manage increases in their distress levels (McNeil and Binder, 1997).

Moreover, it may not be symptoms per se, but skills and confidence in being able to manage symptoms that is most predictive of relapse. The data used in this study (like most typical clinical data) do not include measures of self-efficacy for managing distressing symptoms, or measures of commitment to abstain from high-risk behaviors when under distress. Our analyses were limited by the types of data that are perhaps most commonly collected in PTSD treatment programs as part of a standard assessment protocol. Inclusion of measures of self-

efficacy to maintain sobriety and remain abstinent from high-risk behaviors at the commencement of treatment might yield important information about which subsets of patients continue to place themselves at risk after treatment.

Finally, the results from this study underscore the importance of evaluating high-risk behaviors in this patient population, given the high rates of each of the problem behaviors occurring in a very short period after discharge from a highly structured and controlled therapeutic environment. This may speak to the difficulty in changing established behavior patterns such as violence, suicide attempts, and substance use. However, it may also be indicative of the difficulty current treatment programs may have in targeting these problems, especially in patients who pose substantial risk. For individuals with chronic treatment-resistant PTSD, it would seem that reducing high-risk behaviors should indeed be a major goal of treatment.

There are several limitations to this study. First, it is unclear to what extent these findings generalize to other PTSD samples in and outside of the VA system of care. For example, patients in our sample might present with more chronic and resistant PTSD syndromes that are associated with and driven by different behavioral, demographic, and psychological factors. Therefore, different models may have emerged if tested in different samples. In addition, no therapeutic contextual variables were included in the analysis to assess the impact of treatment factors or the therapeutic milieu on outcomes. Patients' response to treatment, including perhaps their perception of support among treatment staff, might influence behavioral outcomes. Finally, we have not considered variables in postdischarge treatment that might be predictive of patient outcomes.

## CONCLUSION

High-risk behaviors persist among veterans with PTSD despite treatment completion. We have presented models that emerged from empirically driven analyses using signal detection methods in hopes of identifying subgroups of patients at highest risk for negative behavioral outcomes following treatment. However, ROC analysis aimed at identifying high-risk patients did not yield models that were replicable beyond the level of identifying recent behavior as the most robust predictor of future high-risk behaviors. This may suggest that current standard psychometric measures are poor predictors of future behavior. Future research is needed to determine the extent to which empirically based prediction models can be useful in terms of sensitivity and ability to replicate findings, which could then be used to tailor interventions and follow-up procedures to minimize negative outcomes.

## ACKNOWLEDGMENTS

The authors thank Helena Kraemer for her consultation regarding data interpretation.

## REFERENCES

Beck AT, Steer RA, Brown GK (1996) *Beck Depression Inventory* (2nd ed, Manual). San Antonio (TX): Psychological Corp.

Beckham JC, Feldman ME, Kirby AC, Hertzberg MA, Moore SD (1997) Interpersonal violence and its correlates in Vietnam veterans with chronic posttraumatic stress disorder. *J Clin Psychol.* 53:859–869.

Bonin MF, Norton GR, Asmundson GJG, Dicurzio S, Pidlubney S (2000) Drinking away the hurt: The nature and prevalence of PTSD in substance abuse patients attending a community-based treatment program. *J Behav Ther.* 31:55–66.

Chief Medical Director's Special Committee on PTSD (1991) *Program Guide to Specialized Inpatient Post-Traumatic Stress Disorder Units (SIPUs).* Washington DC: Department of Veterans Affairs.

First MB, Spitzer RL, Gibbon M, Williams JBW (1995) *Structured Clinical Interview for DSM-IV Axis I Disorders—Patient Edition (SCID-I/D, Version 2.0).* New York: Biometrics Research Department, New York Psychiatric Institute.

Fontana A, Rosenheck R (1997). *Outcome Monitoring of VA Specialized Intensive PTSD Programs: FY 1996 Report.* West Haven (CT): Northeast Program Evaluation Center.

Fontana A, Rosenheck R (1996) Improving the efficiency of outpatient treatment for posttraumatic stress disorder. *Admin Policy Mental Health.* 23:197–210.

Fontana A, Rosenheck R (1995) An etiological model of attempted suicide among Vietnam theater veterans. *J Nerv Ment Dis.* 183:377–383.

Hammarberg M, Silver SM (1994) Outcome of treatment for post-traumatic stress disorder in a primary care unit serving Vietnam veterans. *J Trauma Stress.* 7:195–216.

Hiley-Young B, Blake DD, Abueg FR, Rozynko V, Gusman FD (1995) Warzone violence in Vietnam: An examination of premiliatry, military and postmilitary factors in PTSD in-patients. *J Trauma Stress.* 8:125–141.

Johnson DR, Rosenheck R, Fontana A, Lubin H, Southwick S, Charney D (1996) Outcome of intensive inpatient treatment for combat-related posttraumatic stress disorder. *Am J Psychiatry.* 153:771–777.

Keane TM, Caddell JM, Taylor KL (1988) Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: Three studies in reliability and validity. *J Consult Clin Psychol.* 56:85–90.

Keane TM, Wolfe J, Taylor KL (1987) Post-traumatic stress disorder: Evidence for diagnostic validity and methods of psychological assessment. *J Clin Psychol.* 43:32–43.

Kiernan M, Kraemer HC, Winkleby MA, King AC, Taylor CB (2001) Do logistic regression and signal detection identify different subgroups at risk? Implications for the design of tailored interventions. *Psychol Methods.* 6:35–48.

Kraemer HC (1992) *Evaluating Medical Tests: Objective and Quantitative Guidelines.* Newbury Park: Sage.

Kulka RA, Schlenger WE, Fairbank JA, Hough RL, Jordan BK, Marmar R, Weiss D (1990) *Trauma and the Vietnam War Generation: Report of Findings From the National Vietnam Readjustment Study.* New York: Brunner/Mazel.

McFall M, Fontana A, Raskind M, Rosenheck R (1999) Analysis of violent behavior in Vietnam combat veteran psychiatric inpatients with posttraumatic stress disorder. *J Trauma Stress.* 12:501–517.

McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, Pettinati H, Argeriou M (1992) The 5th edition of the addiction severity index. *J Subst Abuse Treatment.* 9:199–213.

McNeil DE, Binder RL (1997) The impact of hospitalization on clinical assessments of suicide risk. *Psychiatr Serv.* 48:204–208.

Perconte ST, Griger ML (1991) Comparison of successful, unsuccessful and relapsed Vietnam veterans treated for posttraumatic stress disorder. *J Nerv Ment Dis.* 179:558–562.

Rosenheck R, Fontana A (2001) Impact of efforts to reduce inpatient costs on clinical effectiveness: Treatment of posttraumatic stress disorder in the Department of Veterans Affairs. *Med Care.* 39:168–180.

Savarese VW, Suvak MK, King LA, King DW (2001) Relationships among alcohol use, hyperarousal and marital abuse and violence in Vietnam veterans. *J Trauma Stress.* 14:717–732.